PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Severity of COVID-19 and adverse long-term outcomes: a retrospective cohort study based on a US electronic health record database
AUTHORS	Jovanoski, Nick; Chen, Xin; Becker, Ursula; Zalocusky, Kelly; Chawla, Devika; Tsai, Larry; Borm, Michelle; Neighbors, Margaret; Yau, Vincent

VERSION 1 – REVIEW

REVIEWER	de Figueiredo, Alexandre Universidade Federal da Paraiba, Department of Health Promotion
REVIEW RETURNED	26-Sep-2021

GENERAL COMMENTS	The article is well written and presents a relevant topic. The methodology is well described and the sample is large. The analyzes were well executed and the discussion is consistent with the results. The article also presents the limitations of the methodology used.

REVIEWER	Villalba, Julian Massachusetts General Hospital, Pathology
REVIEW RETURNED	04-Oct-2021

GENERAL COMMENTS	The authors present a very interesting manuscript with a unique large sample that tries to understand a very important question on COVID-19: the presence and prevalence of long-term outcomes of patients affected by the disease. The occurrence and increased risk of development of sequela in COVID-19 has been hypothesized by many authors, but data is scarce, and most cohorts (except few e.g. PMID: 33836148) are relatively small. The large cohort in this study has limitations, but those are described in the manuscript. Studies derived from EHR datasets have limitations, but their large samples can provide insights for future well-designed studies, and therefore I think these data is valuable and should be published in your journal. However, I have a series of questions that I want the authors to address before this paper is published. Comments: -The abbreviation of ECMO is wrong and should get corrected (page 8 -Methods). The right term is Extracorporeal membrane oxygenation. -The authors mentioned that the enrolling period was February 2020 to December 2020. However, in the Methods (page 7 – Section: Modeling and statistical analysis) when they mention that the Date of diagnosis was also binned into months in 2020 (pre-

	April, April, May, June, July), they exclude August, September, October and December. Could you please explain how did you bin those months? Or were patients from those months excluded? -In page 11, in the following sentence: "..indicating that a diagnosis in one category did not necessarily lead to a diagnosis in another.." I would suggest changing the word "indicating" for "suggesting" your analysis does not have the ability to indicate causality. Therefore, this is only a hypothesis. -As the authors mention the specificity of the effect of COVID-19 in the increased long-term outcomes is unclear, as ICU survivors usually develop post-hospitalization sequelae. Therefore, the authors should include this (one sentence) in the abstract conclusion and in the section of strengths and limitations of the study. This study is not designed to identify causality, especially because there is no control group and also due to its retrospective nature. Further studies are needed to answer these questions. -Besides older age, Caucasian race, and non-Hispanic ethnicity, Supplemental figure 1 also suggests that COVID diagnosis in 04/2020 predict new cancer diagnosis >90–≤180 days after COVID-19 diagnosis or hospital discharge. Could the Authors comment on this? This may suggest some bias in the sample used for this study. The authors should comment on this in the discussion of the manuscript. Questions for the authors: -Regarding the code B97.29 (other coronavirus as the cause of diseases classified elsewhere), is it possible that a patient may have been infected with a seasonal coronavirus (e.g.. HCoV-NL63, HCoV-229E, HCoV-OC43 and HCoV-HKU1) instead of SARS-CoV-2, or not? -How was the list of long-term outcomes selected? Did the selection involve a medical team? What was the rationale used to select them? I am asking this, because in some outcomes the pathophysiology may not be related to SARS-CoV-2 infection (e.g. Asthma, COPD, Influenza), therefore someone may think that they would not appear as a sequela of the disease. There is not much clinical and pathophysiological rationale to select some of those outcomes. However, I understand the authors may want to explore if COVID-19 may predispose to the occurrence of these outcomes for reasons that may be unknow. -Could the authors state if there are significant differences in the baseline characteristics among the different subcohort of patients described in Table 1?. -Could the authors also show (could be in the supplemental appendix or main manuscript) if there were statistically significant differences in the prevalence of the different outcomes described in table 2 among the different subcohorts? I can see the relative risks do not overlap in some of these conditions, but I am interested in seeing differences in the frequencies. Strengths -Large cohort of patients with COVID-19 Limitations -Retrospective nature of the study. -As the study was performed from a large EHR database, there may be inherent bias of data collection due to this data collection strategy. HER database-based studies are usually characterized by missing data or confounders as it is usually restricted to examining conditions captured by ICD-10 codes, and therefore the

	data must be interpreted with caution. However, the authors have candidly pointed out this limitation in their manuscript.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Dr. Alexandre de Figueiredo, Universidade Federal da Paraiba
Comments to the Author:
The article is well written and presents a relevant topic. The methodology is well described and the sample is large. The analyses were well executed and the discussion is consistent with the results. The article also presents the limitations of the methodology used.

We thank the reviewer for this comment.

Reviewer: 2
Dr. Julian Villalba, Massachusetts General Hospital
Comments to the Author:
The authors present a very interesting manuscript with a unique large sample that tries to understand a very important question on COVID-19: the presence and prevalence of long-term outcomes of patients affected by the disease. The occurrence and increased risk of development of sequela in COVID-19 has been hypothesized by many authors, but data is scarce, and most cohorts (except few e.g. PMID: 33836148) are relatively small. The large cohort in this study has limitations, but those are described in the manuscript. Studies derived from EHR datasets have limitations, but their large samples can provide insights for future well-designed studies, and therefore I think these data is valuable and should be published in your journal. However, I have a series of questions that I want the authors to address before this paper is published.

Comments:
-The abbreviation of ECMO is wrong and should get corrected (page 8 -Methods). The right term is Extracorporeal membrane oxygenation.

We thank the reviewer for pointing out this error, and have corrected this in the revised manuscript.

-The authors mentioned that the enrolling period was February 2020 to December 2020. However, in the Methods (page 7 – Section: Modeling and statistical analysis) when they mention that the Date of diagnosis was also binned into months in 2020 (pre-April, April, May, June, July), they exclude August, September, October and December. Could you please explain how did you bin those months? Or were patients from those months excluded?

This is a valid point and we thank the reviewer for raising it. The reason that these months are binned until July is that each patient was required to have a minimum of 180 days follow-up. This is stated in the first paragraph of the 'Patients and Study Design' section, and we have now added this to the 'Modeling and statistical analysis' section to help clarify: "Date of diagnosis was also binned into months in 2020 (pre-April, April, May, June, July; allowing for≥180 days follow-up until 31 December 2020 at the latest)." We have also updated the diagnosis dates in the abstract ('Participants' section) for clarity. Finally, we have made some adjustments to Figure 1 to better explain the relationship between the index dates (20 Feb–4 July 2020) and the latest possible date of follow-up (31 December 2020).

-In page 11, in the following sentence: "..indicating that a diagnosis in one category did not

necessarily lead to a diagnosis in another..” I would suggest changing the word “indicating” for “suggesting” your analysis does not have the ability to indicate causality. Therefore, this is only a hypothesis.

We have updated this sentence as suggested by the reviewer.

-As the authors mention the specificity of the effect of COVID-19 in the increased long-term outcomes is unclear, as ICU survivors usually develop post-hospitalization sequelae. Therefore, the authors should include this (one sentence) in the abstract conclusion and in the section of strengths and limitations of the study. This study is not designed to identify causality, especially because there is no control group and also due to its retrospective nature. Further studies are needed to answer these questions.

We have modified the abstract conclusions to highlight this limitation, as suggested by the reviewer. A statement highlighting this limitation is already included in the ‘Strengths and limitations’ section; however, we have modified this to more adequately cover the specific points raised by the reviewer: “The main limitation of this retrospective study is that we use treatment setting as a proxy for COVID-19 severity, and therefore it is difficult to tease out effects specific to the treatment setting (e.g., invasive ventilation) from the underlying COVID-19 severity; any differences that exist between cohorts could bias the results, and as all potential confounders may not be controlled for, the results do not indicate causality.”

-Besides older age, Caucasian race, and non-Hispanic ethnicity, Supplemental figure 1 also suggests that COVID diagnosis in 04/2020 predict new cancer diagnosis >90–≤180 days after COVID-19 diagnosis or hospital discharge. Could the Authors comment on this? This may suggest some bias in the sample used for this study. The authors should comment on this in the discussion of the manuscript.

We thank the reviewer for this point. In this sensitivity analysis, the exposure of interest was hospitalization status while the outcome was diagnosis with cancer, and thus we included monthly indicator covariates in our models to control for the confounding effect that month-specific factors may have on the effects of interest (e.g., hospital capacity changes across time). This model was not designed to examine the effects of calendar month on cancer diagnosis, which would require that we adjust for a completely different set of confounding variables to estimate a valid association between calendar month and cancer diagnosis. For instance, we would not expect
obesity, Charlson comorbidity index, insurance type, gender, or ethnicity to be associated with calendar month, and thus, none of these variables should be in the model as confounders of a calendar month vs. cancer diagnosis model. Thus, we feel that interpreting p-values of confounding variables included in our models as possible indicators of bias is difficult to justify using epidemiologic theory.
Additionally, the sensitivity analysis is purely descriptive and not adjusted for multiple testing; while we do see a significant point estimate for Model B, we do not see the same result in Model A. Nevertheless, the effect of 04/2020 on cancer diagnosis may be driven by factors such as an improvement of hospital capacity (in April compared to other months). Patients who may not have been able to attend a healthcare facility to test for cancer in the other months of the pandemic, could in April attend healthcare facilities to test for cancer. However, an increase in oncology hospital services in one month vs. other months would not be expected to bias the results of this study in a differential manner. In the presence of non-differential bias, one would expect point estimates to be biased conservatively.

Questions for the authors:

-Regarding the code B97.29 (other coronavirus as the cause of diseases classified elsewhere), is it possible that a patient may have been infected with a seasonal coronavirus (e.g.. HCoV-NL63, HCoV-229E, HCoV-OC43 and HCoV-HKU1) instead of SARS-CoV-2, or not?

We agree that while the code B97.29 includes other coronaviruses, it is extremely unlikely that there is a meaningful number of non-SARS-CoV-2 coronavirus-infected patients in our study. It is well established in the literature (e.g., "Human coronavirus circulation in the United States 2014-2017", https://pubmed.ncbi.nlm.nih.gov/29427907/) that non-SARS-CoV-2 coronaviruses largely circulate in the winter months, with percentage positive peaks testing from December to March each year at low proportions. ("117 laboratories reported 854,575 HCoV tests; 2.2% were positive for HCoV-OC43, 1.0% for HCoV-NL63, 0.8% for HCoV-229E, and 0.6% for HCoV-HKU1"). Given that the vast majority (>85%) of patients in our study were diagnosed from April to July, when the official U07.1 COVID-19 diagnosis code was introduced, we consider it unlikely that there was a meaningful proportion of seasonal coronavirus infections in our study. Nonetheless, we have added the following sentence to the limitations section of our manuscript to reflect the possibility: "The B97.29 diagnosis code includes other coronaviruses in addition to SARS-CoV-2 and may therefore be a potential limitation of our study; however, the majority of our COVID-19 cohort (>85%) was diagnosed from April to July using the official U07.1 diagnosis code that is specific to COVID-19, meaning it is unlikely that a substantial number of infections, if any, were from other coronaviruses."

-How was the list of long-term outcomes selected? Did the selection involve a medical team? What was the rationale used to select them? I am asking this, because in some outcomes the pathophysiology may not be related to SARS-CoV-2 infection (e.g. Asthma, COPD, Influenza), therefore someone may think that they would not appear as a sequela of the disease. There is not much clinical and pathophysiological rationale to select some of those outcomes. However, I understand the authors may want to explore if COVID-19 may predispose to the occurrence of these outcomes for reasons that may be unknow.

We took an inclusive approach to identifying potential sequelae because ICD code-based diagnoses are not always accurate, and because there was, and still is, uncertainty about what should be classified as a complication of COVID-19. The reviewer is correct that we do not know if/why COVID-19 would lead to higher risk of asthma, COPD, or influenza. But even in the absence of a hypothesized mechanism, it is important to explore if there is an association with these common causes of the non-specific signs and symptoms that do appear to occur fairly frequently post-COVID. We have clarified this in the 'Modeling and statistical analysis' section of the methods: "LTOs were selected to capture a broad range of potential sequelae, even if there was no strong clinical or pathological rationale for their choice, given the absence of sufficient clinical data regarding established complications associated with COVID-19."

-Could the authors state if there are significant differences in the baseline characteristics among the different subcohort of patients described in Table 1?.

This is an interesting point. Looking at the mix of patients across cohorts in Table 1 shows that differences in baseline characteristics do exist in patients across the cohorts. In fact, the inclusion of these baseline characteristics as controls in our models rests on the assumption that there is/may be differences in patients across cohorts. Their inclusion should, therefore, control for the confounding effect that they may have on the effects of interest. Concerning whether the differences are statistically significantly would not change the interpretation of the results.
As to whether or not we should run significance tests for all the covariates and exposure categories, we feel that this risks confusing the message of the table. In this table alone, there are 37 rows of variables (including subcategories among the 9 main variables) multiplied by 6 columns of variables

(exposure categories), for a total of 185 pairwise p-values that would be reported in the table if we were to compare each covariate for differences in outpatients vs. the remaining exposure categories (ER, ICU, etc). Now, if we were to change the reference point from outpatients vs. other exposure categories to ER patients vs. other exposure categories, we would generate even more p-values as well.
A simple overall p-value assessing for any statistical difference would most definitely return a positive signal, while individual pairwise p-values would be quite numerous and likely distract the reader from the main analysis of the manuscript (while also incurring heavy multiple-testing penalties).
As such, we believe that this analysis would not add much to the current manuscript, where we have performed a more suitable modeling analysis where confounding covariates are adjusted for in our main model.

-Could the authors also show (could be in the supplemental appendix or main manuscript) if there were statistically significant differences in the prevalence of the different outcomes described in table 2 among the different subcohorts? I can see the relative risks do not overlap in some of these conditions, but I am interested in seeing differences in the frequencies.

This is a very interesting point and we thank the reviewer for raising it. However, unlike the results from our model, this type of comparison would be unadjusted and therefore less meaningful than the main results from the modeling study. As such, it may be beyond the scope of the current manuscript, but we agree it is interesting, and could therefore form the basis of a future manuscript.

Strengths
-Large cohort of patients with COVID-19

Limitations
-Retrospective nature of the study.
-As the study was performed from a large EHR database, there may be inherent bias of data collection due to this data collection strategy. HER database-based studies are usually characterized by missing data or confounders as it is usually restricted to examining conditions captured by ICD-10 codes, and therefore the data must be interpreted with caution. However, the authors have candidly pointed out this limitation in their manuscript.

We thank the reviewer for noting that the limitations of our analysis are well described.

VERSION 2 – REVIEW

REVIEWER	Villalba, Julian Massachusetts General Hospital, Pathology
REVIEW RETURNED	08-Nov-2021

GENERAL COMMENTS	Thank you for giving me the opportunity to review the second revision of this manuscript. I thank the authors for addressing the questions posed in the prior revision. I still have a series of comments that would help improving the quality of the manuscript: Comments: -I like that the authors changed the title of the manuscript, as how it was mentioned in the prior revision, this study was not designed to identify causality and it is retrospective in nature. -In the abstract, the authors state as main outcome: "Incidence of new clinical conditions after COVID-19 diagnosis or hospital discharge and the potential effect of disease severity on their risk occurrence". The authors should change the words "disease

	severity" by "treatment setting". If desired they coul include in a parenthesis (as a proxy for disease severity). As discussed in the prior revision, the authors cannot assess with certainty the clinical severity of disease without having more clinical data. Aus the authors noted there are significant confounders when assessing severity by treatment setting, therefore this change is important to make the conclusions explicit to the readers. -In the abstract, in the section Design, "disease severity" should be also changed by "treatment setting" -Similarly, in the last paragraph of the introduction, the authors should change COVID severity in the sentence "... and to understand the role COVID-19 severity plays in the manifestation of these outcomes..." for "treatment setting (as a proxy for disease severity)" -I also suggest changing the same words in the rest of the text, although the authors can include the phrase (as a proxy for disease severity).

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2
Dr. Julian Villalba, Massachusetts General Hospital

Comments to the Author:
Thank you for giving me the opportunity to review the second revision of this manuscript. I thank the authors for addressing the questions posed in the prior revision. I still have a series of comments that would help improving the quality of the manuscript:

Comments:

-I like that the authors changed the title of the manuscript, as how it was mentioned in the prior revision, this study was not designed to identify causality and it is retrospective in nature.